Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet

Cameron, Karen; Hagedorn, Birgit; Dieser, Markus; Christner, Brent C.; Choquette, Kyla; Sletten, Ronald; Crump, Byron; Kellogg, Colleen; Junge, Karen

Published in:
Environmental Microbiology

DOI:
10.1111/1462-2920.12446

Publication date:
2014

Citation for published version (APA):
Cameron, K., Hagedorn, B., Dieser, M., Christner, B. C., Choquette, K., Sletten, R., Crump, B., Kellogg, C., & Junge, K. (2014). Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet. *Environmental Microbiology*, *17*(3), 594-609. https://doi.org/10.1111/1462-2920.12446

Diversity and potential sources of microbiota associated with snow on western portions of the Greenland Ice Sheet

Karen A. Cameron[1*], Birgit Hagedorn[2], Markus Dieser[3**], Brent C. Christner[3], Kyla Choquette[2], Ronald Sletten[4], Byron Crump[5], Colleen Kellogg[5], Karen Junge[1]

[1] Applied Physics Laboratory, University of Washington, Seattle, WA

[2] Environment and Natural Resources Institute, University of Alaska Anchorage, Anchorage, AK

[3] Department of Biological Sciences, Louisiana State University, Baton Rouge, LA

[4] Earth and Space Sciences, University of Washington, Seattle, WA

[5] Earth Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR

Author for correspondence: Karen A. Cameron, Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Box 355640, Seattle, WA 98105-6698. +45 91 33 35 94. cameronk@uw.edu

Running title: Greenland Ice Sheet snow microbial diversity

Current addresses:

* Geological Survey of Denmark and Greenland, Department of Geochemistry, Øster Voldgade 10, 1350 Copenhagen K, Denmark

** Center for Biofilm Engineering, Montana State University, Bozeman, MT

Summary

Snow overlays the majority of the Greenland Ice Sheet (GrIS). However, there is very little information available on the microbiological assemblages that are associated with this vast and climate-sensitive landscape. In this study, the structure and diversity of snow microbial assemblages from two regions of the western GrIS ice-margin were investigated through the sequencing of small subunit rRNA genes. The origins of the microbiota were investigated by examining correlations to molecular data obtained from marine, soil, freshwater and atmospheric environments and geochemical analytes measured in the snow. Snow was found to contain a diverse assemblage of bacteria (*Alphaproteobacteria*, *Betaproteobacteria* and *Gammaproteobacteria*) and eukarya (*Alveolata*, *Fungi*, *Stramenopiles* and *Chloroplastida*). Phylotypes related to archaeal *Thaumarchaeota* and *Euryarchaeota* phyla were also identified. The snow microbial assemblages were more similar to communities characterized in soil than to those documented in marine ecosystems. Despite this, the chemical composition of snow samples was consistent with a marine contribution, and strong correlations existed between bacterial beta diversity and the concentration of Na^+ and Cl^-. These results suggest that surface snow from western regions of Greenland contains exogenous microbiota that were likely aerosolized from more distant soil sources, transported in the atmosphere, and co-precipitated with the snow.

Introduction

Snows' transient nature, seasonal distribution and physical properties drive its widespread influences on climate, hydrology and ecosystem functioning (Jones, 1999, Serreze *et al.*, 2006, Vavrus, 2007). Arctic, Antarctic and Alpine snow has been found to harbor microbial communities (for example; Felip *et al.*, 1995, Thomas and Duval, 1995, Carpenter *et al.*, 2000, Segawa *et al.*, 2005, Bachy *et al.*, 2011, Harding *et al.*, 2011, Hell *et al.*, 2013, Møller *et al.*, 2013), with snow from Svalbard glaciers containing microbial abundances of $2 – 8 \times 10^4$ cells ml^{-1} (Amato *et al.*, 2007, Irvine-Fynn *et al.*, 2012). Arctic snow diversity studies have revealed microbial assemblages dominated by bacterial *Proteobacteria, Bacteroidetes, Actinobacteria, Firmicutes, Acidobacteria* and cyanobacteria, and eukaryotic *Stramenopiles*, *Dikarya* and *Alveolata*, using small subunit ribosomal RNA (rRNA) gene sequencing techniques (Amato *et al.*, 2007, Larose *et al.*, 2010, Bachy *et al.*, 2011, Harding *et al.*, 2011, Hell *et al.*, 2013, Møller *et al.*, 2013). Snow-associated microbial communities have been found to be metabolically active under *in situ* conditions in Alpine, Antarctic and Sub-Arctic locations (Felip *et*

al., 1995, Carpenter *et al.*, 2000, Larsen *et al.*, 2007, Lopatina *et al.*, 2013), and snow microbial community production has been suggested to be nutritionally important to supraglacial, subglacial and ice-marginal environments (Hodson *et al.*, 2005, Wynn *et al.*, 2007, Hodson *et al.*, 2008, Schutte *et al.*, 2009, Telling *et al.*, 2011). Nevertheless, data on the composition, biogeography, origins and metabolic activity of snow communities is sparse. Furthermore, there have been no molecular-based studies to date describing the microbial diversity in snow from the vast (1.7 million km^2; Weidick 1995) and climate-sensitive landscape of the Greenland Ice Sheet (GrIS).

Snowpack ecosystems are believed to be seeded through the aeolian transportation of biota from local (Larose *et al.*, 2010, Harding *et al.*, 2011) as well as more distant sources (Harding *et al.*, 2011). The mobilization of soil particulates from terrestrial surfaces via wind, and the production of marine aerosols during bubble dispersal at water-air interfaces, are important mechanisms for the transportation of microorganisms across local, regional and intercontinental distances (Finlay and Fenchel, 2004, Aller *et al.*, 2005, Rousseau *et al.*, 2005, Smith *et al.*, 2013). Once aerosolized, biotic particles may influence meteorological processes by acting as nuclei for ice crystal formation within clouds (Lohmann and Feichter, 2005, Möhler *et al.*, 2007, Pratt *et al.*, 2009). Over time frames of days to weeks (Burrows *et al*., 2009), bioaersols are transported in the atmosphere and are eventually returned to the Earth's surface through gravitational deposition or by precipitation, where they may immigrate and integrate into ecological processes in the receiving environment (Sattler *et al.*, 2001).

The GrIS possesses a largely ubiquitous snow cover that is free of major geographical features beyond the ice marginal regions. This expanse of snow has a dynamic turnover, with an estimated snowfall of ~ 600 $km^3\ yr^{-1}$ (data from 1979 – 2005; Fettweis, 2007) and an estimated surface run-off from snow and ice melt of 300 $km^3\ yr^{-1}$ (data from 1979 – 2006; Fettweis, 2007) to 400 $km^3\ yr^{-1}$ (data from 2010; Bamber *et al.*, 2012). Snow accumulation is highest within central regions of the ice sheet and melt water generation is larger within the ablation zones (Fettweis, 2007). The GrIS is sensitive to the climate, and consequently, freshwater fluxes from the GrIS have been increasing annually for over two decades, due to warmer surface air temperatures (Hanna *et al.*, 2008, Bamber *et al.*, 2012), and since 1985, the area of snow melt has tended to expand each year (Comiso, 2006). Several extreme melt events have occurred within the last

decade (Comiso, 2006, Tedesco, 2007), the most notable of which occurred on July 12th 2012, where 98.6 % of the GrIS was reported to experience surface melt, as opposed to 43.7 % four days earlier (Nghiem *et al.*, 2012). Due to Greenland's topography and geographical location, the microbiology and ecology of GrIS surface snow environments may differ to those of previously studied surface snows from glacial and sea ice environments (for example Hell *et al.*, 2013, Møller *et al.*, 2013). Snow accumulating on the surface of the GrIS is eventually transformed into melt water, sublimated, compacted into glacial ice or redistributed through wind transportation (reviewed in; Hodson *et al.*, 2008, Larose *et al.*, 2013b). Therefore, an understanding of the microorganisms that are co-deposited with snow provides valuable information about the pool of species that serve as inoculum to all icy habitats associated with the GrIS (Hodson *et al.*, 2005, Wynn *et al.*, 2007, Hodson *et al.*, 2008, Schutte *et al.*, 2009).

In this study, the structure, diversity and abundance of bacterial, archaeal and eukaryotic assemblages from North-West (NW) and South-West (SW) regions of the GrIS (Figure 1) were investigated through Illumina sequencing of the small subunit (16S or 18S) rRNA gene. The specific aims of this study were to characterize the microbial composition of the snow, examine spatial bacterial variability in two regions of the GrIS, and infer the origins of these transient biota through comparison with data from likely source environments (i.e., marine water, sea ice, soil, freshwater and air) using correlation analyses of snow bacterial assemblages and chemical profiles. Together, these studies aimed to improve our understanding of the origins, diversity and potential downstream significance of western GrIS surface snow microbial assemblages.

Experimental Procedures

Study sites

To test the spatial variability of snow biota, snow was sampled along two, three point transects on the surface of the GrIS, located ~ 1200 km apart in the NW (early June 2011; NW.1, NW.2, NW.3) and SW (late May 2012; SW.1, SW.2, SW.3) regions of the GrIS (Figure 1; Table 1). NW transect samples were obtained within the Northern part of the Pituffik Peninsula. Snow was sampled 9.5 km (NW.1), 3.5 km (NW.2) and 1.6 km (NW.3) from the ice margin, and the most westerly point (NW.3) was located 9 km from open fjord water, 17 km from the marine waters of Baffin Bay, and 14 km

from the settlement of Thule. Weather observations recorded within the Thule region since 1971 (Global Historical Climatology Network Database (GHCND) ID: GLW00017605; data available through National Oceanic & Atmospheric Administration) show a mean air temperature of - 11.7 ± 3 °C, with the warmest temperatures recorded in July (5.3 ± 3 °C) and the lowest temperatures recorded in February (- 26.0 ± 5 °C). Annual trends show an increase in air temperature since 1978 (0.04 °C yr^{-1}). The mean precipitation is ~ 124 mm yr^{-1} water equivalent, of which ~ 40 % falls during summer (June, July, August). Points within the SW transect were sampled 53 km (SW.1), 28 km (SW.2) and 2.5 km (SW.1) from the ice margin. The most westerly point of the SW transect (SW.3) lies 50 km from open fjord water, 172 km from the marine waters of the Davis Strait, and 43km from the settlement of Kangerlussuaq. Between 1948 and 2003, the mean annual air temperature within this region was -5.1 °C, with a mean June to August temperature of 9.8 °C (Aelby and Fritz, 2009). The mean annual precipitation was 158 mm yr^{-1}, and the mean calculated evaporation was 300 mm yr^{-1} (Aelby and Fritz, 2009), indicating that this area has a negative water balance. Transect samples were obtained within the ablation zone, however, the most easterly points (NW.1 and SW.1) were sampled within a region that is believed to retain multiyear snow. Two additional snow samples were opportunistically obtained from the NW GrIS region; one from the edge of the ice sheet (NW.IM), and one within the snow transect area in August (NW.4). Furthermore, two freshwater supraglacial lakes, which were at the SW.2 and SW.3 snow transect sites (SW.2.lake, SW.3.lake respectively), were sampled for comparison.

Sampling

All sampling sites were accessed by helicopter, with the exception of site NW.IM, which was accessed from the margin on foot. Snow and supraglacial lake water were sampled for chemical analysis by collecting samples in 4.5 L Whirl-Pak bags at sites that were approximately 10 m from microbiological sampling points (see Table 1). Snow samples were thawed at ~ 25°C and stored at 4 °C in closed bags until analysis at the University of Anchorage Alaska. Snow samples for microbial abundance analysis were collected into sterile 50 ml centrifuge tubes and melted at room temperature. Samples were fixed by adding a 0.22 µm filtered formaldehyde solution to a final concentration of 2 % v/v. The samples were stored at -20 °C until processed at the University of Washington. For the extraction of microbial DNA, approximately 15 kg of snow from the top 30 cm of the snowpack was placed into autoclaved bags using a snow shovel

cleaned with 10 % v/v HCl and wiped with 70 % v/v ethanol. Approximately 15 L of supraglacial lake water was collected into HCl (10 % v/v) cleaned carboys that were rinsed with 0.22 µm filtered deionized water. Snow samples were melted at room temperature within the sampling bags for up to 36 h, after which they were processed immediately. Each sample was filtered through a 0.22 µm Sterivex filter (Millipore, MA, USA), which was then filled with DNA storage buffer (40 mM EDTA, 50 mM Tris HCl, 0.73M sucrose), and stored at -20 °C until DNA extractions were performed at the University of Washington.

Water chemistry

The electrical conductivity and pH of melted snow was measured using a WTW Multi 3430 multiparameter meter (WTW, Weilheim, Germany). Snow and supraglacial lake samples were passed through 0.2 µm filters prior to chemical analysis. Samples collected for cation analysis were acidified to 1 % v/v using HNO_3 (Ultrex-grade) before analysis on an inductively coupled plasma mass spectrometer (ICP MS 7500c; Agilent Technologies, CA, USA). Anion analysis was performed using an ion chromatograph (Dionex IC 500, Dionex, and ICS 5000, Thermo Scientific, MA, USA). An international NIST standard SRM 1643e was used to verify external calibration standards. The limit of detection was calculated with 3.3 times the standard deviation of the calibration regression y-intercept, and the limit of quantitation was defined as 3 times the limit of detection. Accuracy of analysis was < 5 % for concentrations above the reporting limit, and up to 20 % for concentrations below the limit of quantification.

Microbial abundance analysis

Bacterial enumeration was performed by epifluorescent microscopy using a DAPI nucleic acid stain (Invitrogen, NY, USA) alongside an acridine orange counter stain (Invitrogen) on a Zeiss AxioScope A1 microscope, as described by Collins *et al.* (2008).

DNA extraction, amplification and amplicon sequencing

All DNA extractions and reagent preparations for PCR amplification were performed in a laminar flow hood, wiped with ethanol and irradiated with germicidal UV. Filter pipet tips and DNA- and RNAase free certified plasticware was used

throughout. Sterivex filter units were opened and internal filter membranes were removed using a flame-heated razor blade. DNA extractions were performed using PowerWater DNA Isolation Kits (Mo Bio Laboratories Inc., CA, USA). Amplification of the bacterial 16S rRNA gene V4 hypervariable region was achieved using primers detailed by Caporaso *et al.* (F515/R806; 2011). To obtain an overview of archaeal and eukaryotic diversity, PCRs were performed using pooled DNA extracts from NW.1, NW.2 and NW.3 sites (named NW.1-3), and SW.1, SW.2 and SW.3 sites (named SW.1-3). The V6 hypervariable region of archaeal 16S rRNA genes was targeted using V6-major and V6-minor primer sets, as described by Huber *et al.* (958arcF/1048arcR-major/1048arcR-minor; 2007). The eukaryotic 18S rRNA V9 hypervariable region was targeted using euk1 and euk2 primer sets, designed by Amaral-Zettler *et al.* (F1380/F1389/R1510; 2009). All forward and reverse primers were modified to include a unique eight-nucleotide barcode. PCR reaction mixtures contained 1X PCR Gold buffer (Applied Biosystems, CA, USA), 2.5 mM $MgCl_2$, 200 µM of each deoxyribonucleotide triphosphates (Invitrogen, NY, USA), 0.3 µM of each primer, 2.5 U of AmpliTaq Gold - LD Polymerase (Applied Biosystems) and up to 10 pg of the extracted DNA. The reaction volume was adjusted to a total of 50 µl with ultrapure DNase/RNase free water. Thermal cycles consisted of an initial denaturation of 9 min at 95 °C, followed by 43 cycles of 94 °C for 30 sec, 55 °C for 60 sec and 72 °C for 60 sec, and a final extension of 7 min at 72 °C.

To prepare amplicons for sequencing, PCR products were cleaned using QIAquick PCR Purification Kits (Qiagen, Hilden, Germany). Cleaned amplicons were combined into three pools, based on their bacterial, eukaryotic or archaeal origin. The pooled samples contained 250 ng of DNA and were processed with the TruSeq RNA and DNA Sample Preparation Kit (Illumina Inc., CA, USA). Sequencing was conducted on an Illumina MiSeq system (running MCS v.1.2.3 and RTA 1.14.23 software; Illumina, CA, USA).

Sequence read analysis

Forward orientated demultiplexed sequences were quality-filtered and processed using a QIIME processing platform (Caporaso *et al.*, 2010b). QIIME standard operating procedure quality filters were used throughout. V-REVCOMP was used to identify bacterial and archaeal forward orientated sequences (Hartmann *et al.*, 2011). Bacterial and archaeal operational taxonomic units (OTUs) were defined as sequences that possessed ≥ 97% identity, which were clustered

using a reference based UCLUST algorithm against a Greengenes (GG) reference library (Desantis *et al.*, 2006, Edgar, 2010). Bacterial and archaeal sequences were aligned using PyNAST, and taxonomic classifications were assigned by training the Ribosomal Database Project (RDP) classifier (Wang *et al.*, 2007, Caporaso *et al.*, 2010a) to use the February 2011 GG taxonomic dataset. Secondary taxonomic classifications were performed using NCBI BLASTn (Zheng *et al.,* 2000). Eukaryotic OTUs with ≥ 97 % identity were selected using a *de novo* UCLUST OTU clustering method, which allows for the generation of OTU clusters based on sequence identities within the dataset, rather than generating OTU clusters against a database of reference sequences. Eukaryotic taxonomic assignments were made using the Silva 104 reference database (Quast *et al.*, 2013). OTUs that were not taxonomically classified under the Silva 104 reference database were discarded. ChimeraSlayer was used to identify chimeric sequences within all sequence profiles (Haas *et al.*, 2011) and QIIME was used to remove chimeric sequences and singletons. Sequence profiles were rarefied to the number of sequences of the smallest sample sequence output within each pool. CatchAll was used to calculate parametric alpha diversity (Bunge, 2011). Bray-Curtis resemblance, cluster analysis, non-metric multi-dimensional scaling (MDS), analysis of similarity (ANOSIM), contributions of variables to similarity (SIMPER) and multi-variant environmental correlation analysis (BIO-ENV) were calculated from OTU matrices using PRIMER-E version 6 (Plymouth, UK). Amplicon datasets are available at The European Bioinformatics Institute under study accession number PRJEB4904 (www.ebi.ac.uk).

Biogeographical analysis

All snow sequences were compiled alongside 16S rRNA gene sequence profiles from 11 previously reported studies of snow, slush, supraglacial ice, marine surface water, sea ice, periglacial soil, periglacial lake and air environments (Table 2), as well as to the supraglacial lake assemblages investigated within this current study (SW.2.lake, SW.3.lake). Sequences were processed and analyzed using the same methodologies detailed for bacterial sequence read analysis.

Results

Major ion chemical analysis

Chemical analysis of GrIS snow and supraglacial lake samples revealed that NW GrIS samples had higher concentrations of Na^+, Cl^- and SO_4^{2-} ions in comparison to samples obtained from the SW GrIS region (NW GrIS mean Na^+ concentration

was 6.0 fold higher than SW GrIS mean Na^+ concentration; NW GrIS mean Cl^- concentration was 3.8 fold higher than SW GrIS mean Cl^- concentration; NW GrIS mean SO_4^{2-} concentration was 1.6 fold higher than SW GrIS mean SO_4^{2-} concentration; Table 1, Supplemental Figure 1). Within the NW GrIS snow transect, concentrations of Na^+ and Cl^- were found to be highest closest to the ice sheet margin and decreased in concentration further inland (Table 1). The elemental composition of major cations and anions indicate that NW GrIS snow samples had the greatest similarity to seawater (Supplemental Figure 1). The pH at all sites ranged between 5.1 and 5.6, and EC values ranged between 1.7 to 3.7 $\mu S\ cm^{-1}$ (Table 1).

Microbial abundance

Snow microbial abundance from the three NW snow transect sites, NW.1, NW.2 and NW.3, were calculated as $4.7 \times 10^2 \pm 2.8 \times 10^3$ cells ml^{-1}, $2.5 \times 10^2 \pm 2.2 \times 10^3$ cells ml^{-1} and $4.1 \times 10^2 \pm 1.8 \times 10^3$ cells ml^{-1} respectively. However, due to the low microbial abundance, it was necessary to filter large volumes of snow, and cells may have been masked by sediment particles. Hence, our cell abundance data probably underestimate actual cell concentrations. The mean microbial abundance of NW GrIS sampled snow was $3.8 \times 10^2 \pm 1.1 \times 10^4$ cells ml^{-1}. With the exception of the SW.2 snow sample, which had an estimated abundance of $2.6 \times 10^4 \pm 3.6 \times 10^4$ cells ml^{-1}, microbial abundance from SW region snow samples could not be measured due to combined complications from low abundance and the masking effects of sediments.

Bacterial structure, diversity and biogeochemical correlations

PCR amplifications of 16S rRNA genes were successful for all snow and supraglacial lake samples, however, 43 thermal cycles were used to achieve this (as guided by the AmpliTaq Gold manual), which will have likely increased PCR associated artifacts (V. Wintzingerode *et al*., 1997). On average, 50.0 ± 26.9 % of sequences per sample were lost in downstream quality filtering steps (Supplemental Table 1). Quality filtered sequence profiles ranged from 2404 reads per sample (SW.3.lake) to 56803 reads per sample (NW.3) (Supplemental Table 1). Following the CatchAll calculation of parametric alpha diversity, species richness estimates of rarefied snow bacterial profiles were found to range between 357 to 1756 OTUs per sample, with the lowest richness calculated at SW.2 and the highest richness calculated at the NW.1 site (Supplemental Table 1). The mean taxon richness of snow sampled from the NW region (1037 ± 532 OTUs per

sample) was estimated to be more than twice as rich as snow sampled from the SW region (423 ± 94 OTUs per sample). Both NW and SW snow sequence profiles has the highest diversity calculated at transect points that were furthest from the ice margin (NW.1; 1756 OTUs per sample, and SW.1; 531 OTUs per sample). No significant trends were observed between diversity estimates and distance from the ice margin (data not shown). Simprof tested cluster analyses of non-transformed OTU profiles revealed that SW.1 and SW.3 profiles were 78 % similar, and NW.1, NW.2, NW.3 and NW.IM samples were more than 28 % similar to each other. Sequences generated from SW.2 had only 9 % similarity to other samples, and the August NW.4 sample had only 11 % similarity to other samples. Overall, NW and SW OTU profiles shared less than 9 % similarity to each other. Pairwise ANOSIM tests of all non-transformed NW and SW GrIS snow OTU profiles revealed that these assemblages were moderately similar ($R = 0.66$, $P = 0.018$, NW GrIS $n = 5$, SW GrIS $n = 3$). However, on exclusion of the outlying SW.2 and NW.IM OTU profiles from the pairwise ANOSIM analysis, the NW and SW OTU profiles were less similar ($R = 0.86$, $P = 0.067$, NW GrIS $n = 4$, SW GrIS $n = 2$), but were not significant at an $\alpha = 0.05$.

Sequences classifying within *Proteobacteria* dominated the bacterial phylotypes in all GrIS snow samples, with the exception of SW.2 (Figure 2). The mean abundance of *Proteobacteria* was 60.4 ± 25.4 %, which predominantly consisted of sequences related to *Beta-* (50.0 %), *Gamma-* (31.7 %) and *Alphaproteobacteria* (17.9 %). When all western GrIS snow sequences were considered as a single compiled sample, *Sphingobacteria* (9.7 %), *Actinobacteria* (8.1 %), *Acidobacteria* (3.9 %), *Bacilli* (3.7 %), *Clostridia* (3.6 %) and *Flavobacteria* (1.0 %) taxa were also major contributors of species richness. Additionally, 52 classes of bacteria were identified that each represented less than 0.5 % of the total abundance (compressed to the < 0.5 % fraction in Figure 2). OTUs that represented ≥ 10 % of the NW snow assemblages, sampled in June, included GG OTU IDs 102382 (most closely related to *Massilia timonae;* NW.3; 10.6 % of total abundance), 165313 (most closely related to *Acinetobacter johnsonii;* NW.2; 10.3 % of total abundance) and 109056 (most closely related to *Pseudomonas sp.;* NW.1; 22.8 %, NW.3; 16.0 % of total abundance). The August NW snow sample, NW.4, was not heavily represented by any of these OTU sequences, however, over half of the sequence profile was represented by three OTUs that were most closely related to GG OTU IDs 39819 (*Acinetobacter ursingii*; 17.0 %), 143746 (*Candidatus Odyssella*; 19.3 %) and 354510 (*Collimonas arenae*; 20.2 %). No single OTU represented more than 10 % of the sequence

profile generated from the NW ice margin sample (NW.IM). Both SW.1 and SW.3 sequence profiles were dominated by GG OTU IDs 20151 (most closely related to *Arcicella aquatica*; SW.1; 35.0 %, SW.3; 19.3 %) and 80104 (most closely related to *Actimicrobium antarcticum*; SW.1; 40.5 %, SW.3; 47.5 %). Amplicons from SW.2 DNA extractions were most heavily represented by sequences that were related to cyanobacteria (52.3 %, of which 48.6 % were most closely related to the *Phormidium pristleyi* GG OTU ID 184822) and *Acidobacteria* (27.9 %, of which 25.9 % were most closely related to the *Granulicella arctica* GG OTU ID 546864). Cyanobacteria related OTUs constituted < 1 % of all other sequence profiles. OTUs (17) related to ice nucleating species of the *Enterobacteriaceae, Xanthomonadaceae and Pseudomonadaceae* (reviewed in Christner *et al*., 2008 and reported in Joly *et al*., 2013) included *Pseudomonas fluorescens* (5 OTUs), *Pseudomonas syringae* (1 OTU), *Pseudoxanthomonas sp.* (4 OTUs), *Xanthomonas sp.* (2 OTUs), *Pantoea agglomerans* (1 OTU) and *Erwinia* (4 OTUs). These targeted OTUs compiled between 0.2 % (SW.1) and 25.5 % (NW.1) of the total rarefied assemblage population (2404 sequences per sample), with a mean abundance of 5.9 ± 10.0 %. OTUs that were found to be most closely related to *Pseudomonas fluorescens* represented 5.7 ± 9.9 % of the total assemblage population. When SIMPER tests of square root transformed Bray-Curtis OTU similarities were used to reveal order level taxonomies that were responsible for the greatest levels of identity between NW and SW snow profiles, over half of the similarities found (54 %) were accounted for by OTUs that were most closely related to *Burkholderiales*, *Pseudomonadales*, *Sphingobacteriales, Actinomycetales* and *Sphingomonadales* orders.

SW GrIS supraglacial lake biota were sequenced and analyzed to generate OTU profiles from an adjacent yet distinct environmental biome. CatchAll species richness calculations of the rarefied sequence profiles were found to be more than twice as rich as the SW snow samples taken within the vicinity (SW.2.lake; 852 OTUs per sample, SW.3.lake; 843 OTUs per sample). Taxonomies that represented more that 10 % of the SW.2.lake OTU profile included *Gammaproteobacteria* (32.4 %), *Alphaproteobacteria* (14.0 %), *Betaproteobacteria* (13.2 %) and *Actinobacteria* (10.4 %), and sequences from the SW.3.lake were strongly represented by *Betaproteobacteria* (35.7 %), *Gammaproteobacteria* (13.5 %) and *Bacilli* (11.5 %; Figure 2). No single OTU from these profiles dominated assemblages by more than 10 %. CLUSTER analysis of SW GrIS supraglacial lake OTU profiles with GrIS snow profiles revealed that there were minimal

resemblances to profiles generated from the adjacent SW GrIS snow (both SW.2.lake and SW.3.lake had < 10 % similarity to SW snow samples), however, stronger similarities were found to the NW snow samples (> 20 % similarity).

To identify correlations between chemical compositions (including Na^+, Cl^-, Mg^{2+}, K^+, Ca^{2+}, F^-, SO_4^{2-}, NO_3^-; see Table 1) and GrIS snow and supraglacial lake bacterial assemblages, a BIO-ENV statistical analysis was performed. Analyses of square-root transformed OTU profiles were found to correlate most strongly to a combined presence of Na^+ and Mg^{2+} concentrations (ρ = 0.856, P = 0.01). Within this analysis, correlations to singular chemical compositions were highest for Na^+ and Cl^- (ρ = 0.852, ρ = 0.718 respectively), followed by Mg^{2+} (ρ = 0.662), and were poor for SO_4^{2-}, K^+, NO_3^- and Ca^{2+} (ρ < 0.2 for all analyses). Within the 50 strongest diversity correlations to multiple chemical combinations, 41 combinations included Na^+ as a factor influencing the strength of the correlation (ρ ranged from 0.613 to 0.856). Correlations where ρ > 0.8 included influences from combinations of Na^+, Cl^-, Mg^{2+}, K^+, and Ca^{2+}, where up to four ions contributed towards the correlations. BIO-ENV analyses of square-root transformed OTU profiles that were segregated into phyla taxonomies were found to correlate with less statistical significance (P = 0.02 – 0.75), however, OTUs in the *Actinobacteria* were found to strongly correlate with Na^+ (ρ = 0.790, P = 0.02) while *Betaproteobacteria* related OTUs correlated with Na^+ and NO_3^- (ρ = 0.729, P = 0.04).

Archaeal diversity

Archaeal diversity was investigated using two archaeal specific primers sets (V6-major and V6-minor) to amplify 16S rRNA genes from pooled DNA extracts of NW and SW snow traverses. After OTUs were quality filtered, V6-major and V6-minor sequences from pooled NW samples had low sequence outputs (70 and 327 sequences per sample respectively; Supplemental Table 2), therefore, a low rarefaction cut-off point of 70 sequences per sample was selected for all samples. CatchAll calculated alpha diversity of each rarefied sample was 32 ± 10 OTUs per sample post-rarefaction, and up to 365 OTUs per sample pre-rarefaction (Supplemental Table 2).

Archaeal assemblages were comprised of *Thaumarchaeota* and *Euryarchaeota* phyla (Figure 3). *Thaumarchaeota* dominated the archaeal sequences obtained (mean composition of all four profiles; 74.3 ± 9.3 %), with the genus

311 *Candidatus Nitrososphaera* representing 55.7 ± 6.8 % of the mean composition for all four profiles. Both NW (NW.1-3)
312 and SW (SW.1-3) assemblages contained sequences that were related to the *Halobacteria* class (mean abundance of V6-
313 major and V6-minor amplicons; NW.1-3; 5.7 ± 4.0 %, SW.1-3; 25.7 ± 14.1 %). V6-major and V6-minor amplicons that
314 were most related to methanogenic organisms comprised an average of 17.1 % of the archaeal NW.1-3 sequence profile,
315 however, amplicons that were related to methanogenic organisms were absent from the SW.1-3 V6-minor amplicon
316 profile and were only minimally identified within the SW.1-3 V6-major sequence profile (1.4 % composition).
317
318 *Eukaryotic diversity*
319 Eukaryotic sequence profiles were generated through the identification of 18S rRNA genes from pooled NW (NW.1-3)
320 and SW (SW.1-3) snow transect DNA extractions, using two primer sets to target the V9 hypervariable region (total
321 number of quality filtered sequences generated per sequenced amplicon ranged from 38,674 to 83,224). After
322 rarefaction of all profiles to 38,674 sequences per sample, the CatchAll calculated alpha diversity of each pooled transect
323 sample, and from both primer sets (euk1 and euk2) was high (NW.1-3; 1134 and 1252 OTUs estimated, SW.1-3; 836 and
324 862 OTUs estimated respectively).
325
326 Both NW and SW snow transect amplicon assemblages were strongly represented by *Alveolata*, *Fungi*, *Stramenopiles*
327 and *Chloroplastida* (Figure 4). *Fungi* related sequences made up 63.4 % and 65.8 % of the NW euk1 and euk2 primer
328 generated amplicons (respectively). Of these euk1 and euk2 primer generated amplicons, 76.6 % and 73.5 % were most
329 closely related to *Basidiomycota* (respectively; data not shown). *Alveolata* related sequences dominated the SW snow
330 profiles (mean OTU representation of euk1 and euk2 profiles; 62.4 %) and were most strongly represented by sequences
331 that were related to species of the order *Gymnodiniphycidae* (mean OTU representation of euk1 and euk2 profiles;
332 98.1 %; data not shown). Sequences that clustered taxonomically at the phylum level but represented less than 0.5 % of
333 the total sequence profile included taxonomies from 18 different phylum groups within *Alveolata*, *Amoebozoa*,
334 *Centrohelida*, *Discoba*, *Holozoa*, *Metamonada*, *Rhizaria*, *Stramenopiles* and *Chloroplastida*.
335
336 *Bacterial biogeographical analysis*

The bacterial sequence data from this study were compared with archived data from a selection of environmental types representing possible sources of atmospheric bioaerosols (see Table 2). Archaeal and eukaryotic biogeographical analyses were not performed due to a lack of available comparative data. Cluster analysis of square-root transformed Bray-Curtis indices of OTU profiles grouped each environmental type into distinct clades. Exceptions to this included bacterial assemblages sampled from frozen freshwater (including snow, slush, supraglacial surface ice and supraglacial lake environments) and marine environments (marine surface water and sea ice samples, termed marine environments herein) which each clustered into their own separate group (Figure 5; cluster analysis results not shown). In addition, bacterial assemblages sequenced from Colorado snow (Bowers *et al.*, 2009) clustered separately from other frozen freshwater assemblages, with one Bowers *et al.* sample clustering alongside sequences obtained from air samples within the same study, and one sample clustering separately (Figure 5; cluster analysis results not shown). Cluster analysis of each environmental clade revealed that OTU profiles obtained within the same study had higher percentages of similarity (up to 75 % similarity) than those obtained from different studies (< 12 % similarity).

To test for correlations between OTU assemblages sampled from different types of environmental origin, pairwise ANOSIM tests of square root transformed Bray-Curtis similarities were calculated. NW and SW GrIS snow assemblages from this current study were calculated to be most similar to soil sampled assemblages (R = 0.475, P = 0.001 and R = 0.720, P = 0.001 respectively; Supplemental Table 3). SIMPER tests of non-transformed Bray-Curtis similarity indices were used to reveal OTUs that were responsible for the greatest levels of identity between environmental groups. Bacterial order taxonomies that contributed over 5 % of similarities between NW GrIS snow and soil bacterial taxonomic profiles included the *Burkholderiales* (18.7 %), *Actinomycetales* (14.8 %), *Sphingomonadales* (11.8 %), *Clostridiales* (6.4 %) and *Lactobacillales* (6.2%). Similarities between SW GrIS snow and soil bacterial order taxonomies were greatest among *Burkholderiales* (49.9 %), *Sphingobacteriales* (23.7 %) and *Acidobacteriales* (11.4 %; Figure 6). When the OTU profiles of all four snow studies (this study, Bowers *et al.,* 2009, Hell *et al.,* 2013, Møller *et al.,* 2013) were considered as a single sample, strong similarities were found to assemblages characterized in air (R = 0.317, P = 0.001; Supplemental Table 3), moderate similarities were found to soil sampled assemblages (R = 0.689, P = 0.001; Supplemental Table 3) and marine sampled assemblages were found to be dissimilar (R = 0.919, P = 0.001; Supplemental Table 3). However, on

considering each snow study separately, only snow assemblages sampled by Bowers *et al*. (2009) were found to be highly similar to those from air, although, due to having few observations, this analysis was not found to be statistically significant (R = 0.334, P = 0.141; Supplemental Table 3). Of the similarities found between GrIS snow and atmospheric aerosols, OTUs that were most closely related to *Pseudomonadales* and *Burkholderiales* taxa accounted for the greatest percentage of similarities to NW profiles (35.1 % and 20.5 % respectively), and *Burkholderiales* contributed towards 85.8 % of the similarities found to SW profiles (Figure 6).

Both NW and SW GrIS snow assemblages were calculated to lack overlap with assemblages generated from marine environments (R = 0.92, P = 0.001 and R = 0.98, P = 0.001 respectively; Supplemental Table 3). Of the similarities that were identified *Pseudomonadales* (30.0 %), *Actinomycetales* (13.9 %), *Sphingomonadales* (11.1 %) and *Burkholderiales* (9.6 %), and *Burkholderiales* (45.5 %), *Sphingobacteriales* (23.8 %) and *Pseudomonadales* (6.3 %) related OTUs were found to contributed the most towards overlap with NW and SW snow profiles respectively (Figure 6).

Discussion

Permanent snow surfaces and annual snowmelt events influence local ecosystem functioning and have widespread ecological, hydrological and climatological impacts (Jones, 1999, Serreze *et al.*, 2006, Vavrus, 2007). The GrIS is the Earth's second largest body of ice, covering 1.7 million km^2, receiving ~ 600 km^3 of snow each year, and supporting both permanent and seasonal snow covers (Fettweis, 2007, Van Den Broeke *et al.*, 2008). However, despite the magnitude and global importance of the GrIS landmass, this study is the first to report on the structure, diversity and likely origins of GrIS surface snow microbiota. These data provide valuable information on this climate sensitive, underexplored niche, and are of relevance when considering the interconnectivity of the GrIS supraglacial environment to downstream englacial, subglacial, periglacial and marine ecosystems.

Surface snow samples from NW and SW regions of the GrIS were found to contain diverse bacterial and eukaryotic assemblages, with species of archaea also present (Figures 2, 3, 4). Bacterial taxonomic richness estimates from rarefied NW and SW GrIS snow assemblages (i.e. 357 – 1756 OTUs per sample; Supplemental Table 1) were similar to those

calculated from supraglacial cryoconite niches (Cameron *et al.*, 2012), and arable and grassland soils (Hughes *et al.*, 2001, Torsvik *et al.*, 2002). All but one of these bacterial assemblages were dominated by *Proteobacteria*, with NW assemblages being largely composed of *Alpha-*, *Beta-* and *Gammaproteobacteria,* and SW assemblages being strongly represented by the *Betaproteobacteria* (Figure 2), similar to the communities of other supraglacial snow, ice and cryoconite environments (e.g. Amato *et al.*, 2007, Larose *et al.*, 2010, Harding *et al.*, 2011, Cameron *et al.*, 2012, Edwards *et al.*, 2013, Hell *et al.*, 2013). The eukaryotic alpha diversities of rarefied NW and SW GrIS snow transect samples were high (i.e. 836 to 1252 OTUs per sample), especially when compared to a previously constructed small clone library of High-Arctic snow eukaryotes by Bachy *et al.* (38 clones; 2011), where only four eukaryotic OTUs were identified. However, the large eukaryotic alpha diversity estimates calculated within this current study may have been influenced by the *de novo* OTU picking methodology that was applied, which allows for OTU clusters to be generated based on similarities to each other, rather than on a framework of reference sequences. In contrast, alpha diversity indices of heavily rarefied archaeal assemblages were low (i.e. 17 to 38 OTUs per sample when rarefied to 70 sequences per sample), with phylotypes originating from *Thaumarchaeota* and *Euryarchaeota* phyla; this is similar to the archaeal community structure of supraglacial cryoconite holes (Cameron *et al.*, 2012, Edwards *et al.*, 2013). Archaeal alpha diversity estimates calculated on OTU profiles prior to rarefaction were up to an order of magnitude higher (SW.1-3; 365 OTUs estimated per sample; Supplemental Table 2), and were found to be similar to diversity estimates of archaeal communities found within the surface waters of the North Sea (Wemheuer *et al.,* 2012). Cyanobacteria were neither a major component of the GrIS surface snow biota analyzed within this current study, nor of Svalbard (Amato *et al.*, 2007, Hell *et al.*, 2013), Canadian Arctic (Harding *et al*., 2011) and Tibetan (Liu *et al.*, 2009) snow communities reported in other studies. An exception to this was noted within the SW.2 GrIS snow bacterial profile, where *Phormidium pristleyi* related OTUs dominated (49 % relative abundance; Figure 2); perhaps as a result of local environmental conditions or biotic sources. Larose *et al.* (2010) similarly found anomalies in the presence of cyanobacteria within Svalbard snow samples, suggesting that there is a necessity to further develop biogeographical datasets, to determine robust ecological patterns between local, regional and inter-landmass snow environments.

Early summer NW and SW GrIS surface snow bacterial compositions were found to cluster by location, with the exception of the SW.2 sample, however, trends in diversity along the transects were not identified. In contrast, studies of GrIS surface ice bacteria, sampled along a transect by Telling *et al.* (2012), found that the relative abundance of nitrogen fixation genes initially increased with distance from the ice margin, implying that community structure and function fluctuate spatially. Additionally, studies of Himalayan and Alaskan snow revealed relationships between altitude and algal abundance and diversity (Yoshimura *et al.*, 1997, Nozomu, 2013). NW and SW GrIS locations had differing cation and anion compositions (Supplemental Table 1), and at both locations concentrations of Na^+ and Cl^- were found to increase with proximity to marine waters. Similarly, maritime influences on snow chemistries have previous been shown to vary with altitude and distance from marine sources within the Terra Nova Bay region of Antarctica (Udisti *et al*., 1999). The unique nature of the August sampled NW GrIS snow biotic profile (NW.4; 11 % similarity to other samples) within this current study hints towards surface snow assemblages either undergoing compositional changes in response to environmental conditions, as outlined in snow studies by Segawa *et al.* (2005), Hell *et al.* (2013) and Larose *et al.* (2013c), or experiencing temporal variability based on the deposition of biotic assemblages alongside snow, similar to the temporal studies of airborne microbial communities by Fierer *et al.* (2008). Bacterial assemblages amplified from SW GrIS snow and adjacent supraglacial lakes lacked similarity. Additionally, neither bacterial nor eukaryotic sequence profiles from SW GrIS snow, sampled in May 2012, had resemblances to the longer-established supraglacial cryoconite communities, sampled within the same GrIS region by Cameron *et al.* (2012) in late August 2008, despite potentially being seeded by similar sources. While this lack of identity to communities from surrounding biomes once again suggests that snow microbiota are heavily influenced by the biotic assemblages present at the time of snow deposition, further temporal, spatial and geochemical investigations are undoubtedly essential to ascertain the factors that drive the structure and diversity of GrIS surface snow microbiota.

When the bacterial assemblages in snow were compared to communities from marine, soil, supraglacial, freshwater and air samples, strong to moderate ANOSIM test resemblances were calculated to profiles of soil bacteria (NW GrIS snow; $R = 0.48$, $P = 0.001$, SW GrIS snow; $R = 0.72$, $P = 0.001$, all four snow studies compiled; $R = 0.69$, $P = 0.001$; Supplemental Table 3), suggesting that the biotic composition of the sampled surface snow, was largely influenced by the translocation

of terrestrially originating microorganisms. Similarities between snow and air sampled assemblages were additionally found (Supplemental Table 3), which was largely driven by the presence of *Pseudomonadales* and *Burkholderiales* orders (Figure 6). OTUs which were found to be closely related to taxa previously described as having ice nucleating activities (reviewed in Christner *et* al., 2008 and reported in Joly *et al*., 2013) contributed 5.9 ± 10.0 % of the total number of snow sequences analyzed, and of these, 5 OTUs that were related to *Pseudomonas fluorescens* contributed 5.7 ± 9.9 %. While the full extent of bacterial taxonomies and strains that possess ice nucleating properties remains unknown (Christner *et al*., 2008), the significant proportion of bacteria with the potential to perform ice nucleation investigated within this current study further suggest the connection between microbial air transport and snow deposition. Lesser resemblances were found between snow and marine sampled assemblages (NW GrIS snow; R = 0.92, P = 0.001, SW GrIS snow; R = 0.98, P = 0.001, all four snow studies compiled; R = 0.92, P = 0.001; Supplemental Table 3). Despite this, chemical analyses of snow samples revealed strong maritime signatures (Supplemental Figure 1), and furthermore, correlations were calculated between OTU profiles and Na^+ and Cl^- chemical compositions (ρ = 0.852, ρ = 0.718 respectively). Interestingly, studies of Canadian Arctic bacterial snow clone libraries by Harding *et al.* (2011), have previously identified the presence of *Glaciecola*, *Colwellia*, *Loktanella* and *Polaribacter* genera, which are typically found in cold oceanic waters.

While the biogeographical analysis used within this current study is a useful tool to speculate on the origins and seeding mechanisms of snow sampled environments, it is important to note the confinements that are imposed due to the limited availability of comparable published sequence data. In addition, methodological discrepancies in sampling and nucleic acid extractions, amplifications and sequencing between studies reduces their comparability. Study bias effects were most apparent among the Bowers *et al*. (2009) snow and air sampled sequences, which were found to cluster independently from other studies, regardless of environmental type (Figure 5), and which swayed the calculated ANOSIM relationship between air and compiled snow assemblages (Supplemental Table 3). Despite this, the clustering of assemblages, generated through other independent investigations, by environmental type, is nonetheless convincing of the effectiveness of this methodology (Figure 5).

While investigations into the biogeochemical activities of snow were not studied, the high diversity of bacterial and eukaryotic assemblages identified lends itself towards the potential for a wide range of metabolic activities to be performed. Snow microbial communities from other geographic locations have been found to be biogeochemically active, with noted roles in carbon and nitrogen cycling (Felip *et al.*, 1995, Larsen *et al.*, 2007, Telling *et al.*, 2012, Larose *et al.*, 2013a). Considering the extent of the GrIS surface snow biome, if active, snow communities contribute noteworthy carbon and nitrogen inputs to global budgets. Despite the low abundance of cells associated with snow sampled from the NW region of Greenland (3.8×10^2 cells ml^{-1}), when considered alongside estimates of GrIS surface runoff in 2010 (~ 400 km^3 yr^{-1}; Bamber *et al.*, 2012), the total cellular abundance associated with GrIS supraglacial runoff entering downstream environments is likely to be in the region of 1.5×10^{20} cells yr^{-1}; with a carbon equivalent of 4.5 ± 1.85 Mg C yr^{-1}, and a nitrogen equivalent of 0.9 ± 0.2 Mg N yr^{-1} [based on mean cellular carbon (30.2 ± 12.3 fg C $cell^{-1}$) and nitrogen (5.8 ± 1.5 fg C $cell^{-1}$) contents of surface coastal bacterial assemblages; Fukuda *et al*., 1998]. Similar calculations, done using the higher cellular abundance of snow sampled from the SW GrIS region (2.6×10^4 cells ml^{-1}), estimated the annual cellular content of GrIS surface meltwater fluxes to be ~ 1.0×10^{22} cells yr^{-1}, with a carbon equivalent of $3.14 \times 10^2 \pm 1.28 \times 10^2$ Mg C yr^{-1}, and a nitrogen equivalent of 60.3 ± 15.6 Mg N yr^{-1}. These assemblies of cells will likely have little impact on the surrounding biologically and nutritionally rich open ocean environments, where biomass equivalents are calculated to be found in 0.8 – 5.6 km^3 of Arctic Ocean waters (based on marine abundance measurements by Bowman *et al.*, 2012). However, microbiota originating from GrIS surface snow may, nonetheless, provide valuable nutritional and genetic resources to biotic niches within more proximal downstream englacial, subglacial, periglacial, estuary and coastline environments.

In summary, the sampled GrIS surface snow environment was found to contain distinct and diverse biotic assemblages, containing bacteria, eukaryotes and archaea; with strong representation from bacterial *Alpha*-, *Beta*- and *Gammaproteobacteria,* and eukaryotic *Alveolata*, *Fungi*, *Stramenopiles* and *Chloroplastida*. Snow biota resembled soil sampled assemblages, suggesting that these environments are predominantly seeded by wind transported terrestrial sources. The fate of these microorganisms could result in their embedment into the GrIS during snow accumulation, or alternatively they could be relocated away from the GrIS surface through melt and wind processes. As a warming

491 climate continues to increase GrIS surface melt rates annually (Hanna *et al.*, 2008, Bamber *et al.*, 2012), understanding
492 the ecological composition and functionality of the snow environment, and deciphering the impacts of biotic processes
493 on downstream environments, is necessary for establishing its biogeochemical role in polar ecology.
494
495 *Acknowledgments*
496 This work was funded by a National Science Foundation grant (NSF - PLR 1023462, 1023233, 1023351), awarded to K.
497 Junge, B. Christner, B. Hagedorn and R. Sletten; B. Christner was also partially supported by funds from NSF 1241161.
498 Fieldwork in Greenland was supported by Polar Field Services (Colorado, USA), and Lu Liu, Zoe Harrold, Laura Quante
499 and Samuel Foredyce provided field assistance. Shelly Carpenter assisted with abundance analysis and logistical support.
500
501 *Conflict of Interest*
502 The authors declare no conflict of interest.
503
504 *References*

505 Aebly, F.A., Fritz, S.C. (2009) Palaeohydrology of Kangerlussuaq (Søndre Strømfjord), West Greenland during the last ~
506 8000 years. The Holocene 19: 91-104.

507 Aller, J.Y., Kuznetsova, M.R., Jahns, C.J., Kemp, P.F. (2005) The sea surface microlayer as a source of viral and bacterial
508 enrichment in marine aerosols. J Aerosol Sci 36: 801-812.

509 Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., Huse, S.M. (2009) A method for studying protistan diversity using
510 massively parallel equencing of v9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE 4: e6372.

511 Amato, P., Hennebelle, R., Magand, O., Sancelme, M., Delort, A-M., Barbante, C*., et al.* (2007) Bacterial characterization
512 of the snow cover at Spitzberg, Svalbard. FEMS Microbiol Ecol 59: 255-264.

513 Bachy, C., López-García, P., Vereshchaka, A., Moreira, D. (2011) Diversity and vertical distribution of microbial
514 eukaryotes in the snow, sea ice and seawater near the North Pole at the end of the polar night. Front Microbiol 2: 106

Bamber, J., van den Broeke, M., Ettema, J., Lenaerts, J., Rignot, E. (2012) Recent large increases in freshwater fluxes from Greenland into the North Atlantic. Geophys Res Lett 39: L19501.

Bowers, R.M., Lauber, C.L., Wiedinmyer, C., Hamady, M., Hallar, A.G., Fall, R*., et al.* (2009) Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. Appl Environ Microbiol 75: 5121-5130.

Bowman, J.S., Rasmussen, S., Blom, N., Deming, J.W., Rysgaard, S., Sicheritz-Ponten, T. (2012) Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. ISME J 6: 11-20.

Bunge, J. (2011) Estimating the number of species with CatchAll. Biocomputing 121-130.

Burrows, S.M., Butler, T., Jöckel, P., Tost, H., Kerkweg, A., Pöschl, U*., et al*. (2009) Bacteria in the global atmosphere – part 2: Modeling of emissions and transport between different ecosystems. Atmos. Chem. Phys. 9: 9281-9297.

Cameron, K.A., Hodson, A.J., Osborn, A.M. (2012) Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. FEMS Microbiol Ecol 82: 254-67.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., Knight, R. (2010a) Pynast: A flexible tool for aligning sequences to a template alignment. Bioinformatics 26: 266-267.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K*., et al.* (2010b) Qiime allows analysis of high-throughput community sequencing data. Nature Methods 7: 335-336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J*., et al.* (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 108: 4516-4522.

Carpenter, E.J., Lin, S., Capone, D.G. (2000) Bacterial activity in South Pole snow. Appl Environ Microbiol 66: 4514-4517.

Christner, B.C., Cai, R., Morris, C.E., McCarter, K.S., Foreman, C.M., Skidmore, M.L*., et al.* (2008) Geographic, seasonal, and precipitation chemistry influence on the abundance and activity of biological ice nucleators in rain and snow. Proc Natl Acad Sci U S A 105: 18854-18859.

Collins, R.E., Carpenter, S.D. and Deming, J.W. (2008) Spatial heterogeneity and temporal dynamics of particles, bacteria, and pEPS in Arctic winter sea ice. J Marine Syst 74: 902-917.

539 Comiso, J.C. (2006) Arctic warming signals from satellite observations. Weather 61: 70-76.

540 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K.*, et al.* (2006) Greengenes, a chimera-checked
541 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72: 5069-5072.

542 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-2461.

543 Edwards, A., Pachebat, J.A., Swain, M., Hegarty, M., Hodson, A.J., Irvine-Fynn, T.D.*, et al.* (2013) A metagenomic
544 snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. Environ Res Lett 8: 035003.

545 Felip, M., Sattler, B., Psenner, R., Catalan, J. (1995) Highly active microbial communities in the ice and snow cover of high
546 mountain lakes. Appl Environ Microbiol 61: 2394-401.

547 Fettweis, X. (2007) Reconstruction of the 1979–2006 Greenland Ice Sheet surface mass balance using the regional
548 climate model MAR. The Cryosphere Discuss 1: 123-168.

549 Fierer, N., Liu, Z., Rodríguez-Hernández, M., Knight, R., Henn, M., Hernandez, M.T. (2008) Short-term temporal variability
550 in airborne bacterial and fungal populations. Appl Environ Microbiol 74: 200-207.

551 Finlay, B.J., Fenchel, T. (2004) Cosmopolitan metapopulations of free-living microbial eukaryotes. Protist 155: 237-244.

552 Fukuda, R., Ogawa, H., Nagata, T., Koike, I. (1998) Direct determination of carbon and nitrogen contents of natural
553 bacterial assemblages in marine environments. Appl Environ Microbiol 64: 3352-3358.

554 Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G.*, et al.* (2011) Chimeric 16S rRNA sequence
555 formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 21: 494-504.

556 Hanna, E., Huybrechts, P., Steffen, K., Cappelen, J., Huff, R., Shuman, C.*, et al.* (2008) Increased runoff from melt from
557 the Greenland Ice Sheet: A response to global warming. J Climate 21: 331-341.

558 Harding, T., Jungblut, A.D., Lovejoy, C., Vincent, W.F. (2011) Microbes in High Arctic snow and implications for the cold
559 biosphere. Appl Environ Microbiol 77: 3234-3243.

560 Hartmann, M., Howes, C.G., Veldre, V., Schneider, S., Vaishampayan, P.A., Yannarell, A.C., Quince, C., Johansson, P.,
561 Björkroth, K.J. and Abarenkov, K. (2011) V-REVCOMP: Automated high‐throughput detection of reverse

562 complementary 16S rRNA gene sequences in large environmental and taxonomic datasets. FEMS Microbiol Letters 319:
563 140-145.

564 Hell, K., Edwards, A., Zarsky, J., Podmirseg, S.M., Girdwood, S., Pachebat, J.A*., et al.* (2013) The dynamic bacterial
565 communities of a melting High Arctic glacier snowpack. ISME J 7: 1814-1826.

566 Hodson, A.J., Mumford, P.N., Kohler, J., Wynn, P.M. (2005) The High Arctic glacial ecosystem: New insights from nutrient
567 budgets. Biogeochem 72: 233-256.

568 Hodson, A.J., Anesio, A.M., Tranter, M., Fountain, A., Osborn, M., Priscu, J*., et al.* (2008) Glacial ecosystems. Ecol
569 Monographs 78: 41-67.

570 Huber, J.A., Mark Welch, D.B., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A*., et al.* (2007) Microbial population
571 structures in the deep marine biosphere. Science 318: 97-100.

572 Hughes, J.B., Hellmann, J.J., Ricketts, T.H., Bohannan, B.J.M. (2001) Counting the uncountable: Statistical approaches to
573 estimating microbial diversity. Appl Environ Microbiol 67: 4399-4406.

574 Irvine-Fynn, T.D.L., Edwards, A., Newton, S., Langford, H., Rassner, S.M., Telling, J*., et al.* (2012) Microbial cell budgets of
575 an Arctic glacier surface quantified using flow cytometry. Environ Microbiol 14: 2998-3012.

576 Joly, M., Attard, E., Sancelme, M., Deguillaume, L., Guilbaud, C., Morris, C.E*., et al.* (2013) Ice nucleation activity of
577 bacteria isolated from cloud water. Atmospheric Environ 70: 392-400.

578 Jones, H.G. (1999) The ecology of snow-covered systems: A brief overview of nutrient cycling and life in the cold.
579 Hydrological Process 13: 2135-2147.

580 Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D*., et al.* (2010) Microbial sequences retrieved
581 from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway. Extremophiles 14: 205-
582 212.

583 Larose, C., Aurélien, D., Timothy, M.V. (2013a) Microbial nitrogen cycling in Arctic snowpacks. Environ Res Lett 8: 035004.

584 Larose, C., Dommergue, A., Vogel, T. (2013b) The dynamic Arctic snow pack: An unexplored environment for microbial
585 diversity and activity. Biol 2: 317-330.

586 Larose, C., Prestat, E., Cecillon, S., Berger, S., Malandain, C., Lyon, D., Ferrari, C., Schneider, D., Dommergue, A. and Vogel,
587 T.M. (2013c) Interactions between snow chemistry, mercury inputs and microbial population dynamics in an arctic
588 snowpack. PloS one 8 e79972.

589 Larsen, K.S., Grogan, P., Jonasson, S., Michelsen, A. (2007) Respiration and microbial dynamics in two subarctic
590 ecosystems during winter and spring thaw: Effects of increased snow depth. Arct Antarct Alp Res 39: 268-276.

591 Liu, Y., Yao, T., Jiao, N., Kang, S., Xu, B., Zeng, Y*., et al.* (2009) Bacterial diversity in the snow over Tibetan plateau glaciers.
592 Extremophiles 13: 411-423.

593 Lohmann, U., Feichter, J. (2005) Global indirect aerosol effects: A review. Atmospheric Chem Phys 5: 715-737.

594 Lopatina, A., Krylenkov, V., Severinov, K. (2013) Activity and bacterial diversity of snow around russian antarctic stations.
595 Res Microbiol 164: 949-958.

596 Möhler, O., DeMott, P.J., Vali, G., Levin, Z. (2007) Microbiol and atmospheric processes: The role of biological particles in
597 cloud physics. Biogeosci 4: 1059-1071.

598 Møller, A.K., Søborg, D.A., Al-Soud, W.A., Sørensen, S.J., Kroer, N. (2013) Bacterial community structure in High-Arctic
599 snow and freshwater as revealed by pyrosequencing of 16S rRNA genes and cultivation. Polar Res 32: 17390

600 Nghiem, S.V., Hall, D.K., Mote, T.L., Tedesco, M., Albert, M.R., Keegan, K*., et al.* (2012) The extreme melt across the
601 Greenland Ice Sheet in 2012. Geophys Res Lett 39: L20502

602 Nozomu, T. (2013) Seasonal and altitudinal variations in snow algal communities on an Alaskan glacier (Gulkana Glacier
603 in the Alaska range). Environ Res Lett 8: 035002.

604 Pratt, K.A., DeMott, P.J., French, J.R., Wang, Z., Westphal, D.L., Heymsfield, A.J., *et al.* (2009) In situ detection of
605 biological particles in cloud ice-crystals. Nature Geosci 2: 398-401.

606 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P*., et al.* (2013) The SILVA ribosomal RNA gene database
607 project: Improved data processing and web-based tools. Nucleic Acids Res 41: D590-D596.

608 Rousseau, D-D., Schevin, P., Duzer, D., Cambon, G., Ferrier, J., Jolly, D*., et al.* (2005) Pollen transport to Southern
609 Greenland: New evidences of a late spring long distance transport. Biogeosci Discuss 2: 829-847.

Sattler, B., Puxbaum, H., Psenner, R. (2001) Bacterial growth in supercooled cloud droplets. Geophys Res Lett 28: 239-242.

Schütte, U.M., Abdo, Z., Bent, S.J., Williams, C.J., Schneider, G.M., Solheim, B.*, et al.* (2009) Bacterial succession in a glacier foreland of the High Arctic. ISME J 3: 1258-68.

Segawa, T., Miyamoto, K., Ushida, K., Agata, K., Okada, N., Kohshima, S. (2005) Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. Appl Environ Microbiol 71: 123-130.

Serreze, M.C., Barrett, A.P., Slater, A.G., Woodgate, R.A., Aagaard, K., Lammers, R.B.*, et al.* (2006) The large-scale freshwater cycle of the Arctic. J Geophys Res: Oceans 111: C11010.

Smith, D.J., Timonen, H.J., Jaffe, D.A., Griffin, D.W., Birmele, M.N., Perry, K.D.*, et al.* (2013) Intercontinental dispersal of bacteria and archaea by transpacific winds. Appl Environ Microbiol 79: 1134-1139.

Tedesco, M. (2007) A new record in 2007 for melting in Greenland. EOS, Trans American Geophys Union 88: 383.

Telling, J., Anesio, A.M., Tranter, M., Irvine-Fynn, T., Hodson, A., Butler, C.*, et al.* (2011) Nitrogen fixation on Arctic glaciers, Svalbard. J Geophys Res: J Geophys Res 116: G03039.

Telling, J., Stibal, M., Anesio, A.M., Tranter, M., Nias, I., Cook, J.*, et al.* (2012) Microbial nitrogen cycling on the Greenland Ice Sheet. Biogeosci 9: 2431-2442.

Thomas, W.H., Duval, B. (1995) Sierra Nevada, California, U.S.A., snow algae: Snow albedo changes, algal-bacterial interrelationships, and ultraviolet radiation effects. Arct Alp Res 27: 389-399.

Torsvik, V., Øvreås, L., Thingstad, T.F. (2002) Prokaryotic diversity--magnitude, dynamics, and controlling factors. Science 296: 1064-1066.

Udisti, R., Becagli, S., Castellano, E., Traversi, R., Vermigli, S., Piccardi, G. (1999) Sea-spray and marine biogenic seasonal contribution to snow composition at Terra Nova Bay, Antarctica. Annals Glaciol 29: 77-83.

van den Broeke, M., Smeets, P., Ettema, J., Munneke, P.K. (2008) Surface radiation balance in the ablation zone of the West Greenland Ice Sheet. J Geophys Res: Atmospheres 113: D13105.

Vavrus, S. (2007) The role of terrestrial snow cover in the climate system. Climate Dynamics 29: 73-88.

V. Wintzingerode, F., Göbel, U.B. and Stackebrandt, E. (1997) Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. 21: 213-229.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261-5267.

Weidick, A. (1995) Greenland. In Satellite image atlas of glaciers of the world. Williams, R.S., Jr, and Ferrigno, J.G. (eds). US Geol. Surv. Prof. Pap. 1386-C, C1–C93.

Wemheuer, B., Wemheuer, F., Daniel, R. (2012) RNA-based assessment of diversity and composition of active archaeal communities in the German Bight. Archaea 2012: 695826

Wynn, P.M., Hodson, A.J., Heaton, T.H.E., Chenery, S.R. (2007) Nitrate production beneath a High Arctic glacier, Svalbard. Chem Geology 244: 88-102.

Yoshimura, Y., Kohshima, S., Ohtani, S. (1997) A community of snow algae on a Himalayan glacier: Change of algal biomass and community structure with altitude. Arct Alp Res 29: 126-137.

Zheng, Z., Schwartz, S., Wagner, L., Miller, W. (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7: 203-14.

Table and Figure Legends

Table 1 - Location and biochemical details of samples.

Table 2 - Details of studies included in biogeographical analysis. n.p. - not published.

Figure 1 - Map of sampling sites. Crosses indicate biotic assemblages that were sampled within this current study. Details of samples can be found within Table 1. Circles indicate the sampling sites of sequencing studies performed by other authors, which were utilized within this current analysis. Details of referenced studies can be found within Table 2. Dotted line represents the Arctic Circle.

Figure 2 - Relative abundance of class level bacterial taxonomies based on PCR amplifications of 16S rRNA gene sequences.

Figure 3 - Relative abundance of genus level archaeal taxonomies based on PCR amplifications of 16S rRNA gene sequences, using V6-major and V6-minor primers sets with pooled extractions from NW and SW snow transect samples.

Figure 4 - Relative abundance of kingdom [phylum / subkingdom] eukaryotic taxonomies based on PCR amplifications of 18S rRNA gene sequences, using euk1 and euk2 primers sets with pooled extractions from NW and SW snow transect samples.

Figure 5 - Multidimensional scaling plot of square-root transformed Bray-Curtis indices of OTU profiles from multiple compiled studies of differing environmental origins. Circles represent frozen freshwater samples. Triangles represent all other environmental sample types. Letters represent the author of each study, as detailed in Table 2. C - NW and C - SW indicates samples from the current study. Stress factor; 0.16.

Figure 6 - SIMPER calculated percentage contribution of order level taxonomies to similarities between A) NW and B) SW bacterial GrIS snow transect assemblages and bacterial communities of other environmental origins. Taxa without > 0.5 % contribution in at least one profile are not shown.

Table 1 - Location and biochemical details of samples

Sample	Sample type	Region	Location	Sample date	pH	EC $\mu S\ cm^{-1}$	Na^+ $\mu g\ L^{-1}$	Cl^- $\mu g\ L^{-1}$	Mg^{2+} $\mu g\ L^{-1}$	K^+ $\mu g\ L^{-1}$	Ca^{2+} $\mu g\ L^{-1}$	F^- $\mu g\ L^{-1}$	SO_4^{2-} $\mu g\ L^{-1}$	NO_3^- $\mu g\ L^{-1}$
NW.1	Snow transect	NW Greenland - Thule	76.44 N, 67.85 W	9 June 2011	5.6	2.4	16.5	-	-	-	14.0	2.5	195.9	72.9
NW.2	Snow transect	NW Greenland - Thule	76.48 N, 68.10 W	9 June 2011	5.2	3.3	157.9	283.3	16.9	7.3	3.6	2.1	184.3	116.2
NW.3	Snow transect	NW Greenland - Thule	76.51 N, 68.18 W	9 June 2011	5.4	3.7	260.2	500.3	26.4	10.3	19.5	2.1	150.5	67.3
NW.4	August snow	NW Greenland - Thule	76.46 N, 67.93 W	31 August 2011	5.4	2.0	39.9	156.4	8.4	4.3	12.4	2.1	94.2	41.5
NW.IM	Ice margin snow	NW Greenland - Thule	76.53 N, 68.19 W	13 June 2011	-	-	468.4	1114.5	131.9	31.0	270.2	298.6	991.9	519.1
SW.1	Snow transect	SW Greenland - Kangerlussuaq	67.01 N, 48.62 W	18 May 2012	5.3	2.4	22.0	138.2	4.9	6.2	22.4	2.9	109.9	126.9
SW.2	Snow transect	SW Greenland - Kangerlussuaq	67.12 N, 49.37 W	18 May 2012	5.1	2.3	26.4	151.4	5.1	10.7	23.3	3.3	139.8	175.9
SW.3	Snow transect	SW Greenland - Kangerlussuaq	67.23 N, 49.89 W	18 May 2012	5.6	1.7	-	16.4	-	-	-	2.3	74.1	74.1
SW.2.lake	Supraglacial lake	SW Greenland - Kangerlussuaq	67.12 N, 49.37 W	18 May 2012	-	-	-	-	-	-	-	-	-	-
SW.3.lake	Supraglacial lake	SW Greenland - Kangerlussuaq	67.23 N, 49.89 W	18 May 2012	5.2	7.3	264.1	488.8	37.6	42.9	72.1	-	378.4	300.9

- indicates data not available

Table 2 - Details of studies included in biogeographical analysis. n.p. - not published.

Environment	Region	Location	Study accession number	Reference	Key
Snow	Colorado, USA	40.5N 108.7W	Qiime database - Study 314 **	Bowers *et al*., 2009	Bw
	Svalbard	78.1N 15.4E	PRJEB1743	Hell *et al*., 2013	H
	NE Greenland	81.6N 16.7W	SRP003408	Møller et al, 2013	M
	NW and SW Greenland	**See Table 1**	**PRJEB4904**	**This study**	**C - NW, C - SW**
Slush	Svalbard	78.1N 15.4E	PRJEB1743	Hell *et al*., 2013	H
Supraglacial Ice	Svalbard	78.1N 15.4E	PRJEB1743	Hell *et al*., 2013	H
Supraglacial Lake	**SW Greenland**	**See Table 1**	**PRJEB4904**	**This study**	**C - SW**
Marine Water	High Arctic	88.7N 158.9W, 88.5N 129.3W, 88.7N 58.5W	SRP006990	Bowman *et al*., 2012	Bm
	NW Passage	78.7N 104.9W	Earth Microbiome Project - Study 723*	n.p. (Plymouth Marine Laboratory)	P
	Canadian Basin	n.a.	SRP018324	n.p. (Zhou)	Z
Sea Ice	High Arctic	89.5N 129.3W, 88.7N 69.8W	SRP006990	Bowman *et al*., 2012	Bm
	NW Passage	78.7N 104.9W	Earth Microbiome Project - Study 723*	n.p. (Plymouth Marine Laboratory)	P
Soil	NW Greenland, Norway, Canadian Arctic	76N 68W, 70N 19E, 63N 68W, 73N 78W, 82N 62W, 79N 90W	SRP017487	Bell *et al*., 2013	Bl
	E Greenland	74.5N 20.5W	Earth Microbiome Project - Study 1034*	n.p. (Gittel)	G
	Svalbard	78.9N 11.8E	SRP002015	Schütte *et al*., 2009	S
Periglacial lake	NE Greenland	81.6N 16.6W	SRP003408	Møller et al, 2013	M
Air	Colorado, USA	40.5N 108.7W	Qiime database - Study 314 **	Bowers *et al*., 2009	Bw
	Antarctica	78.1S 163.8E	PRJEB1657	n.p. (International Centre for Terrestrial Antarctic Research)	I
		*	http://www.microbio.me/emp/		
		**	http://www.microbio.me/qiime/fusebox.psp		

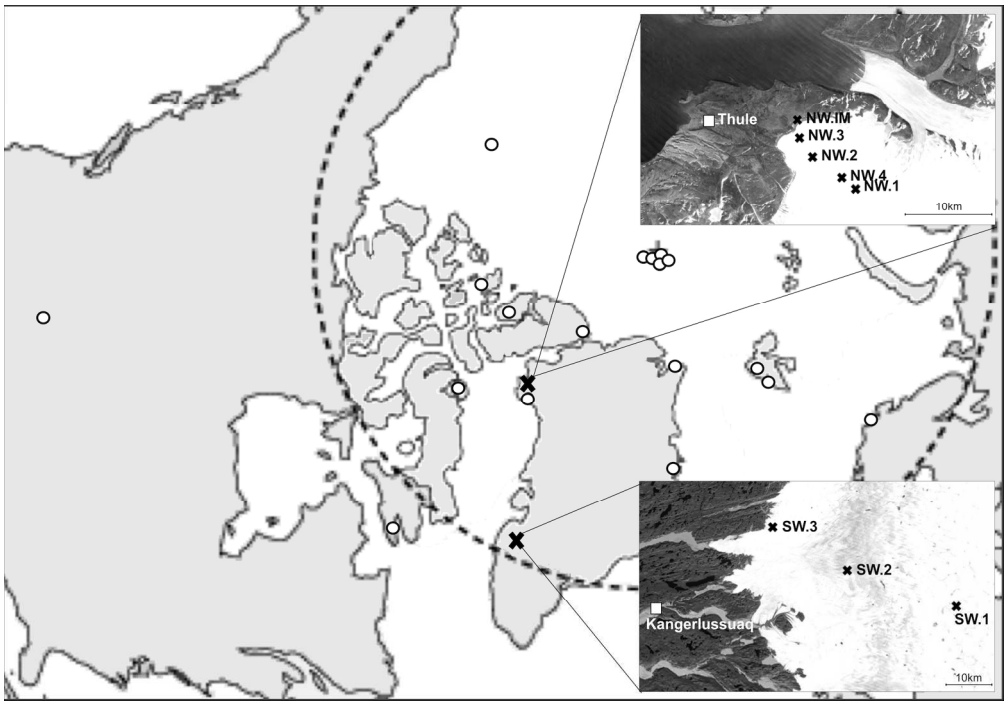


Figure 1 - Map of sampling sites. Crosses indicate biotic assemblages that were sampled within this current study. Details of samples can be found within Table 1. Circles indicate the sampling sites of sequencing studies performed by other authors, which were utilized within this current analysis. Details of referenced studies can be found within Table 2. Dotted line represents the Arctic Circle.
508x353mm (100 x 100 DPI)

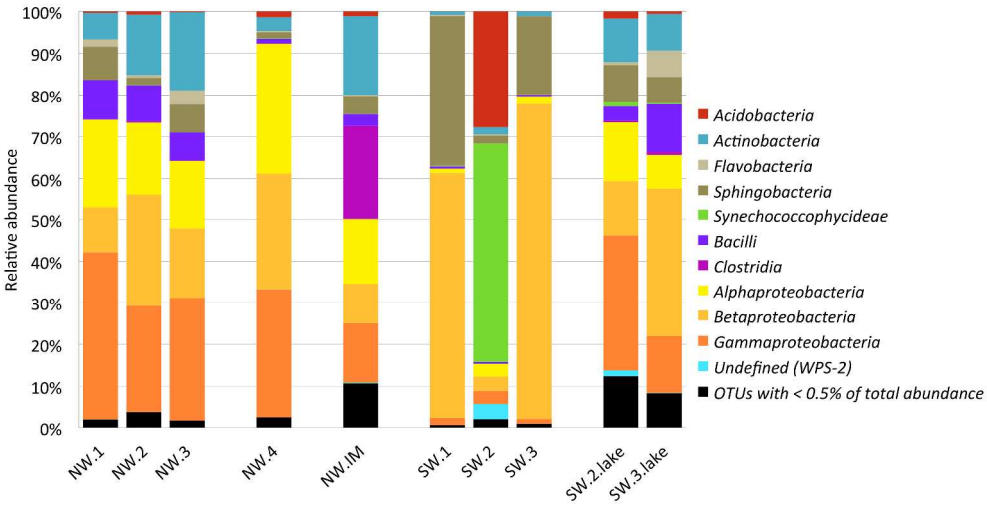


Figure 2 - Relative abundance of class level bacterial taxonomies based on PCR amplifications of 16S rRNA gene sequences.

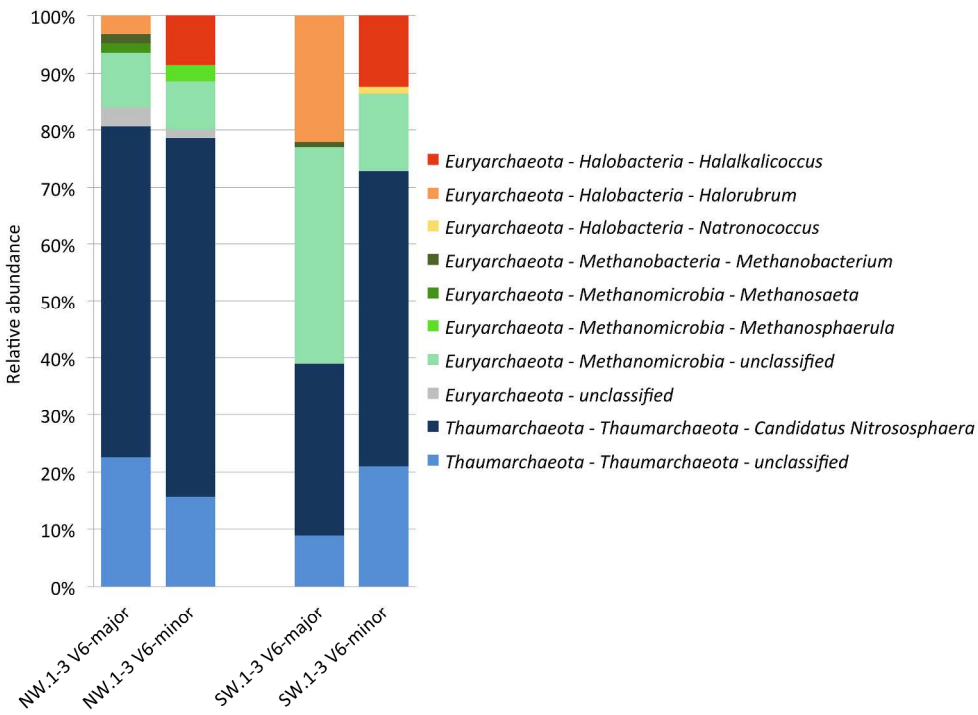


Figure 3 - Relative abundance of genus level archaeal taxonomies based on PCR amplifications of 16S rRNA gene sequences, using V6-major and V6-minor primers sets with pooled extractions from NW and SW snow transect samples.

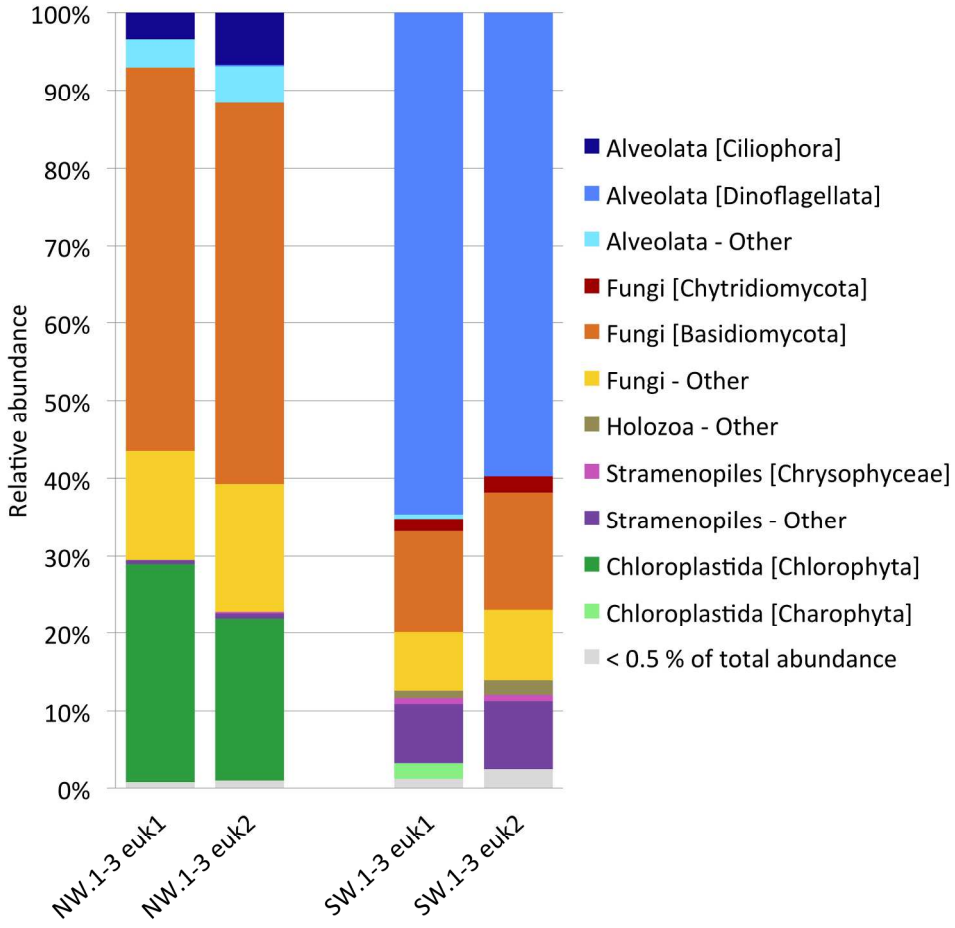


Figure 4 - Relative abundance of kingdom [phylum / subkingdom] eukaryotic taxonomies based on PCR amplifications of 18S rRNA gene sequences, using euk1 and euk2 primers sets with pooled extractions from NW and SW snow transect samples.

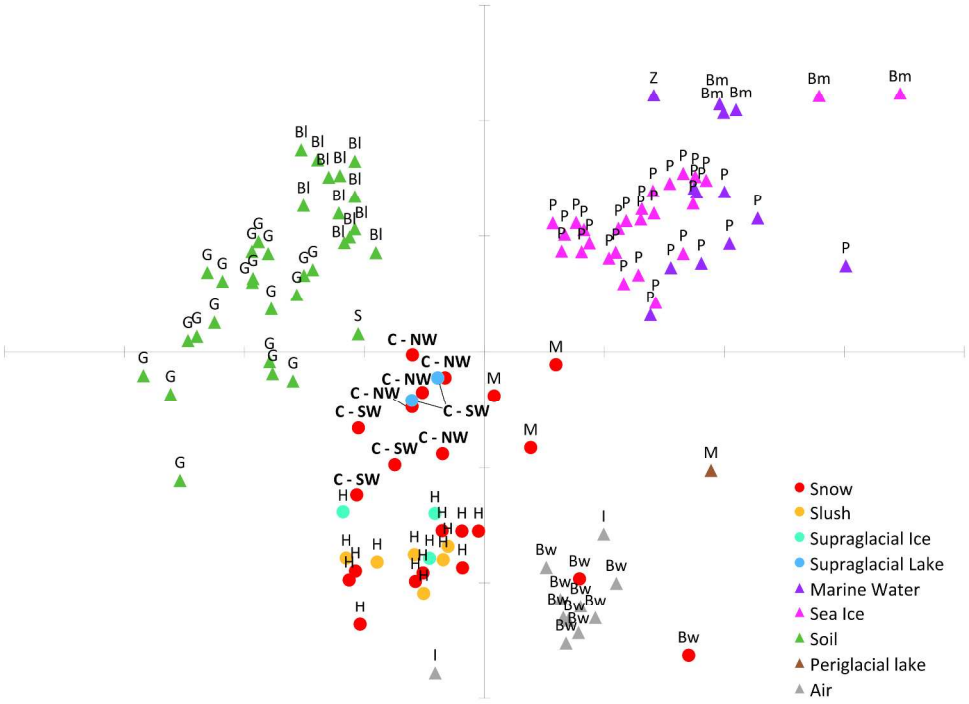


Figure 5 - Multidimensional scaling plot of square-root transformed Bray-Curtis indices of OTU profiles from multiple compiled studies of differing environmental origins. Circles represent frozen freshwater samples. Triangles represent all other environmental sample types. Letters represent the author of each study, as detailed in Table 2. C - NW and C - SW indicates samples from the current study. Stress factor; 0.16.

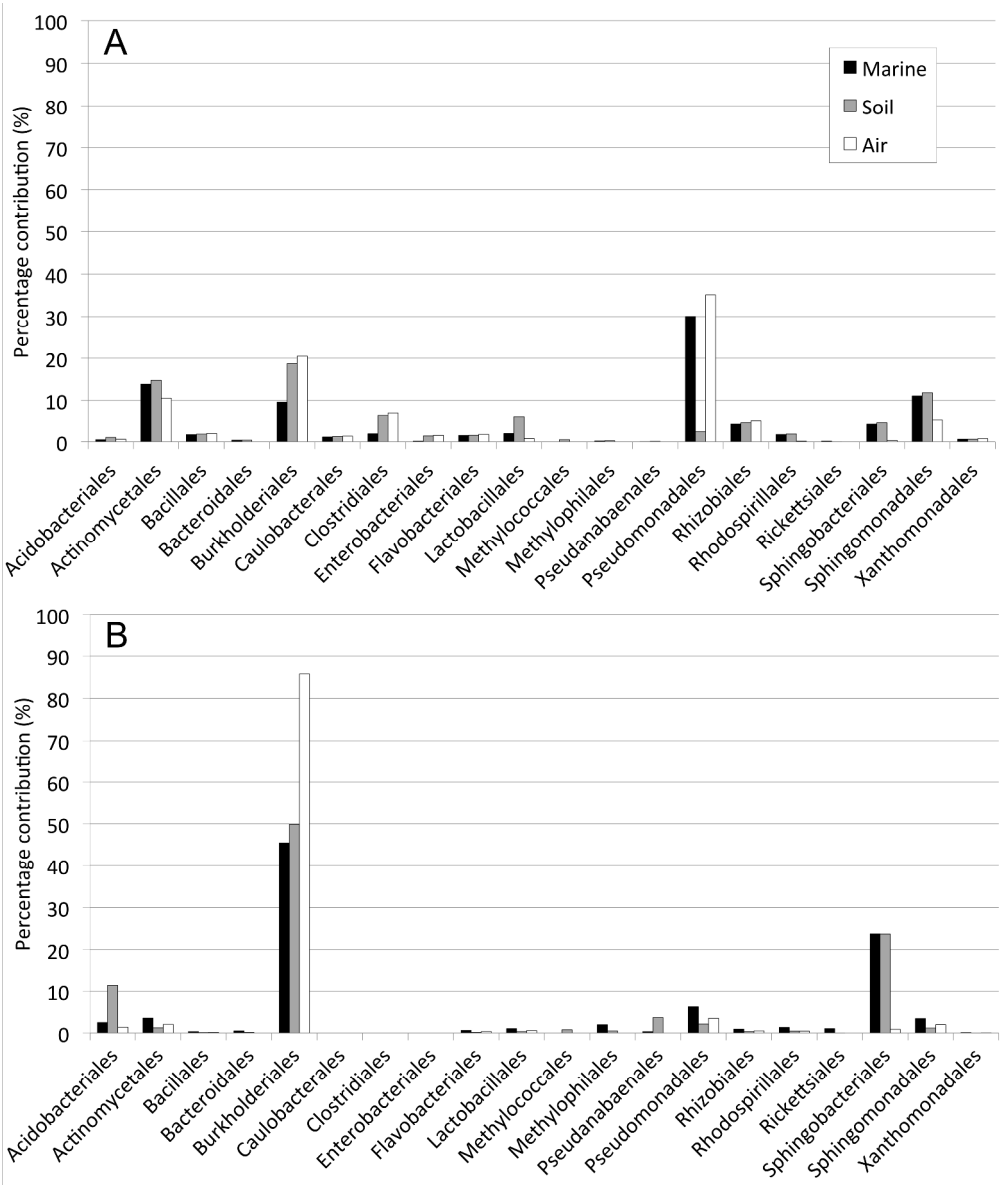


Figure 6 - SIMPER calculated percentage contribution of order level taxonomies to similarities between A) NW and B) SW bacterial GrIS snow transect assemblages and bacterial communities of other environmental origins. Taxa without > 0.5 % contribution in at least one profile are not shown.

Supplemental Table 1

Number of bacterial sequences per sample pre and post quality filtering, and CatchAll diversity estimates pre and post profile rarefaction to 2404 sequences per sample

Amplicon	Sequences per sample pre QF	Sequences per sample post QF	Percentage sequence loss through QF (%)	Diversity estimates pre rarefaction	Diversity estimates post rarefaction
NW.1	53620	35161	34.4	1669.9	1756.1
NW.2	8410	5636	33.0	1209.9	1232.2
NW.3	72559	56803	21.7	1457.6	464.6
NW.4	61404	22268	63.7	1137.1	555.7
NW.IM	20410	8153	60.1	1349.9	1175.6
SW.1	29468	25587	13.2	1235.8	531.9
SW.2	14466	3077	78.7	532.2	357.4
SW.3	19355	12901	33.3	1018.3	382.4
SW.2.lake	32271	11361	64.8	1661.1	852.4
SW.3.lake	77656	2404	96.9	836.9	843.7

Supplemental Table 2

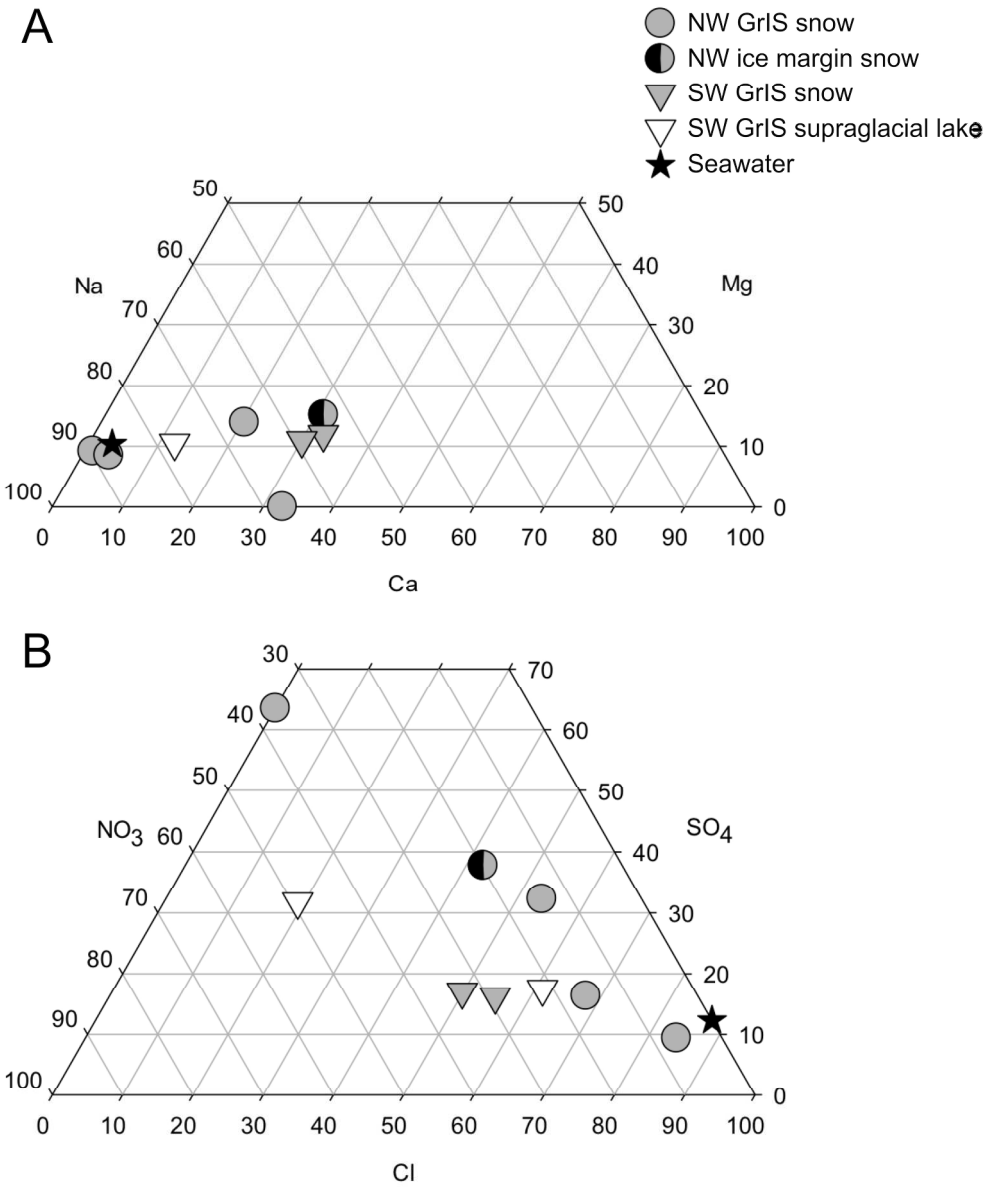
Number of archaeal sequences per sample post quality filtering, and CatchAll diversity estimates pre and post profile rarefaction to 70 sequences per sample

Amplicon	Sequences per sample	Diversity estimates pre rarefaction	Diversity estimates post rarefaction
NW.1-3 V6-major	70	38.3	38.3
NW.1-3 V6-minor	327	170.6	37.5
SW.1-3 V6-major	7638	229	16.5
SW.1-3 V6-minor	16752	365	35.6

Supplemental Table 3

Pairwise ANOSIM tests of square root transformed Bray-Curtis similarities between snow bacterial OTU assemblages and bacterial assemblages sampled from marine (marine water and sea ice), soil and air environments. Compiled snow analyses are highlighted in gray and consisted of OTU assemblages from all five snow locations (4 snow studies). Results with significance levels of ≤ 0.003 are shown in gray text.

		Snow					
	Snow sampling location	**Compiled snow**	**NW GrIS**	**SW GrIS**	**Svalbard**	**NE Greenland**	**Colorado**
	Reference		This Study	This study	Hell *et al.* 2013	Møller *et al*. 2013	Bowers *et al.* 2009
Marine	R	0.919	0.919	0.975	0.989	0.931	0.999
	Significance level	0.001	0.001	0.001	0.001	0.001	0.001
Soil	R	0.689	0.475	0.720	0.856	0.656	0.932
	Significance level	0.001	0.001	0.001	0.001	0.002	0.002
Air	R	0.317	0.860	0.957	0.782	0.714	0.334
	Significance level	0.001	0.002	0.003	0.001	0.008	0.141



Supplemental Figure 1:
Percentage ratio composition of A) anions and B) cations sampled from NW and SW GrIS snow and supraglacial lakes. Symbols plotted as zero indicate data not available. Seawater sample is referenced from Stumm and Morgan (1996).

Supplementary Table and Figure Legends

Supplemental Table 1

Number of bacterial sequences per sample pre and post quality filtering, and CatchAll diversity estimates pre and post profile rarefaction to 2404 sequences per sample

Supplemental Table 2

Number of archaeal sequences per sample post quality filtering, and CatchAll diversity estimates pre and post profile rarefaction to 70 sequences per sample

Supplemental Table 3

Pairwise ANOSIM tests of square root transformed Bray-Curtis similarities between snow bacterial OTU assemblages and bacterial assemblages sampled from marine (marine water and sea ice), soil and air environments. Compiled snow analyses are highlighted in gray and consisted of OTU assemblages from all five snow locations (4 snow studies). Results with significance levels of ≤ 0.003 are shown in gray text.

Supplemental Figure 1

Percentage ratio composition of A) anions and B) cations sampled from NW and SW GrIS snow and supraglacial lakes. Symbols plotted as zero indicate data not available. Seawater sample is referenced from Stumm and Morgan (1996).